Modelling of the Dependence Structure of Regime-Switching Models' Residuals Using Autocopulas

Anna Petričková

Department of Mathematics
Faculty of Civil Engineering
Slovak University of Technology Bratislava
Radlinského 11, 813 68 Bratislava, Slovakia
petrickova@math.sk

Abstract: The autocorrelation function describing linear dependence is not suitable for the description of the residual dependence of regime-switching models. Therefore, we would like to investigate the description of this dependence with a 'k-lag auto-copula', which is a 2-dimensional joint distribution function of the bivariate random vector (Yt , Yt−k) of time lagged values of random variables that generate time series (in the analogy of the autocorrelation function of stationary time series). In this contribution, we will describe the dependence of time lagged residuals of SETAR models by means of copulas, and we will test the independence of these residuals.

Keywords: time series; regime-switching models; residuals; autocopula; product copula; goodness-of-fit tests for autocopulas

1 Introduction

The first models used for modelling economical and financial time series had a linear character (shocks were assumed to be uncorrelated but not necessarily independent and identically distributed - iid). Although many of the models commonly used in empirical finance are linear, the nature of financial data suggests that nonlinear models are more appropriate [5]. Therefore, in recent years, increasing attention has been given to modelling and forecasting economic time series by non-linear models, such as bilinear models, neural networks, regime-switching models, etc. Among other types of non-linear time series models, there are models to represent the changes of variance along time (heteroskedasticity). These models are called autoregressive conditional heteroskedasticity (ARCH) and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models. Here changes in variability are related to, or predicted by, recent past values of the observed series. In this paper we focus on the model SETAR (from the class of regime-switching models).

The autocorrelation function is suitable for the description of the residual dependence only in the case of linear models. So the autocorrelation function is not suitable for the description of the residual dependence of regime-switching models (because these models have nonlinear character).

Therefore we investigate the description of this dependence with '*k*-lag auto-copula', which is a 2-dimensional joint distribution function of the bivariate random vector (Y_t , Y_{t-k}) of time lagged values of random variables that generate time series (in the analogy of the autocorrelation function of linear stationary time series).

First we must test independence in the residuals $\{\hat{e}_t\}$. For our case we use the BDS test. When the BDS test shows residual dependence at a significant level, we use *k*-lag autocopulas for the modelling of these dependence residuals.

The paper is organized as follows. After a general introduction, the theoretical basis of SETAR model, copulas and some tests are described. The paper continues with their application to modelling the dependence of residuals of real time series with auto-copulas.

2 Theoretical Basis

2.1 Model SETAR

In this paper we focus on the class of regime-switching models that are good to interpret and are also very suitable for modeling a large amount of real data. The basic feature of these models is their "control" with one or more variables.

Typical models belonging to this class are TAR models ("**T**hreshold **A**uto**R**egressive"). They form the basis of regime-switching models with regimes determined by observable variables. These models assume that any regime in time *t* can be given by any observed variable q_t (*indicator variable*). Values of q_t are compared with *threshold value c*. In the case of a 2-regime model, the first regime applies if $q_t \le c$, the second if $q_t > c$.

We have the model SETAR when the variable q_t is taken to be a lagged value of the time series itself, that is $q_t = X_{t-d}$ for a certain integer $d > 0$. The resulting model is called a **S**elf-**E**xciting **T**hreshold **A**uto**R**egressive (SETAR) model. For example the 2-regime model SETAR with AR(p) in both regimes has the form

$$X_t = (\phi_{0,1} + \phi_{1,1}X_{t-1} + \cdots + \phi_{p,1}X_{t-p})[1 - \mathbf{1}(X_{t-d} > c)] + \\ + (\phi_{0,2} + \phi_{1,2}X_{t-1} + \cdots + \phi_{p,2}X_{t-p})\mathbf{1}(X_{t-d} > c) + e_t \qquad (1)$$

where $\{e_t\}$ is the strict white noise process with $E[e_t] = 0$, $D[e_t] = \sigma_e^2$ for all $t = 1, ..., n$ and $\mathbf{1}$(A) is the *indicator function* with values $\mathbf{1}$(A) = 1 if the event A occurs and $\mathbf{1}$(A) = 0 otherwise.

In the case of a 3-regime model, we must define 2 constants c_1, c_2 where $-\infty \le c_1 < c_2 \le \infty$. Model SETAR with AR(p) in all regimes has the form

$$X_t = \phi_{0,j} + \phi_{1,j}X_{t-1} + ... + \phi_{p,j}X_{t-p} + e_t \qquad \text{if } c_{j-1} < X_{t-d} \le c_j, \quad j = 1, 2, 3 \qquad (2)$$

For more details see [1], [5].

2.2 The BDS Test

This test was presented in [2] and can be used to test independence in residuals $\{\hat{e}_t\}$. For some $n \in N$ and $\varepsilon > 0$ is the test based on the correlation integral

$$C_{n,\varepsilon} = 2[(T_n - 1)T_n]^{-1} \sum_{m+1 \le \tau < t \le T_n} \sum \mathbf{1}(\|\hat{\mathbf{e}}_{\mathbf{t,n}} - \hat{\mathbf{e}}_{\tau,\mathbf{n}}\| < \varepsilon),$$

where $T_n = T - n + 1$, $\hat{\mathbf{e}}_{\mathbf{t,n}} = (\hat{e}_t, \ldots, \hat{e}_{t+n-1})'$, $\mathbf{1}$(A) is the indicator of the event A, and $\|\cdot\|$ denotes the maximum norm (also known as the Chebyshev norm) in $\Re^d$ (i.e., $\|\mathbf{z}\| = \max_{1 \le i \le d} |z_i|$ for $\mathbf{z} = (z_1, \ldots, z_d)'$). Then the BDS statistic is

$$\Lambda_{BDS} = [(T - m)/V_{n,\varepsilon}]^{1/2} (C_{n,\varepsilon} - C_{1,\varepsilon}^n) \qquad (3)$$

where

$$V_{n,\varepsilon} = 4K_\varepsilon^n + 4(n-1)^2 C_\varepsilon^{2n} - 4n^2 K_\varepsilon C_\varepsilon^{2(n-1)} + 8\sum_{j=1}^{n-1} K_\varepsilon^{n-j} C_\varepsilon^{2j},$$

$$K_\varepsilon = (T - m)^{-3} \sum_{K=m+1}^{T} \sum_{\tau=m+1}^{T} \sum_{t=m+1}^{T} \mathbf{1}(|\hat{e}_K - \hat{e}_\tau| < \varepsilon)\mathbf{1}(|\hat{e}_\tau - \hat{e}_t| < \varepsilon),$$

$$C_\varepsilon = (T - m)^{-2} \sum_{\tau=m+1}^{T} \sum_{t=m+1}^{T} \mathbf{1}(|\hat{e}_\tau - \hat{e}_t| < \varepsilon)$$

and also T is the length of the time series, m is the order of the process AR and n embedding dimension (in our case a lag order of the residuals).

Λ_{BDS} has a N(0,1) asymptotic distribution when $\{e_t\}$ are i.i.d.

When the BDS test at a significant level shows residual dependence, we use k-lag autocopulas for modelling these dependent residuals.

2.3 Copula

2-dimensional copula is a function (see e.g [8])

$$C : [0,1]^2 \rightarrow [0,1] \tag{4}$$

such that

$$C(0, y) = C(x, 0) = 0, \ C(1, y) = y, C(x, 1) = x,$$

for all $x, y \in [0, 1]$ and

$$C(x_1,y_1) + C(x_2, y_2) - C(x_1,y_2) - C(x_2,y_1) \geq 0$$

for all $x_1, x_2, y_1, y_2 \in [0, 1]$ with $x_1 \leq x_2, y_1 \leq y_2$.

The most important applications of 2-dimensional copulas are related to a well known, very convenient alternative of expressing the joint distribution function of 2-dimensional random vectors (X, Y) in the form

$$F(x, y) = C(F_X(x), F_Y(y)), \tag{5}$$

where F_X, F_Y are marginal distribution functions.

Let X, Y be some continuous random variables with joint distribution function $F(x,y)$ and copula C satisfying (5).

Kendall's tau for the random vector (X, Y) is defined (cf. [4]) by

$$\tau(X, Y) = P\{(X - \tilde{X})(Y-\tilde{Y}) > 0\} - P\{(X-\tilde{X})(Y-\tilde{Y}) < 0\}, \tag{6}$$

where $(\tilde{X}, \tilde{Y})$ is an independent copy of (X, Y).

It is well know that (cf. [4])

$$\tau(X, Y) = 4 \int\int_{[0,1]^2} C(u, v) dC(u, v) - 1. \tag{7}$$

2.3.1 Archimedean Class of Copulas

There are many classes of copulas, but in this paper we will use only copulas from the Archimedean class.

Copula C belongs to the Archimedean class if (see e.g. [7], [8], [4])

$$C_\phi(u,v) = \phi^{(-1)}(\phi(u) + \phi(v)) \quad \text{for } u, v \in (0,1],$$

where ϕ: $(0, 1] \rightarrow [0,\infty)$ is a convex, decreasing function (satisfying $\phi(1) = 0$) that is called a generator of the copula C_ϕ, and $\phi^{(-1)} : [0,\infty) \rightarrow [0, 1]$ is given by

$$\phi^{(-1)}(x) = \sup\{\, t \in (0,1] \mid \phi(t) \geq x \} = \begin{cases} \phi^{-1}(x) & x < \phi(0^+) \\ 0 & else \end{cases}.$$

2.3.2 Characteristics of some Archimedean Copulas

As a generator uniquely determines an Archimedean copula, different choices of generators yield many families of copulas that consequently, in addition to the form of the generator, differ in the number and the range of parameters. We summarize some basic facts related to the most important one-parameter families of Archimedean copulas (see e.g. [4]). Note that Clayton and Gumbel copulas model only positive dependence (measured by the Kendall's τ), while Frank covers the whole range [-1, 1].

The following useful relation for Archimedean copulas are presented in [4]

$$\tau = 1 + 4\int_0^1 \frac{\phi(t)}{\phi'(t)}\, dt\,, \tag{8}$$

Gumbel family

Generator $\phi(t) = (-\ln t)^\theta$, where $\theta \geq 1$,

$C_\theta(u, v) = e^{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}}$,

Kendall's $\tau = \dfrac{\theta - 1}{\theta}$.

Strict Clayton family (Kimeldorf and Sampson)

Generator $\phi(t) = \dfrac{t^{-\theta} - 1}{\theta}$, where $\theta > 0$,

$C_\theta(u, v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta}$ and $C_0(u, v) = \Pi = u\,v$,

Kendall's $\tau = \dfrac{\theta}{\theta + 2}$.

Frank family

Generator $\phi(t) = -\ln\left(\dfrac{e^{-\theta t} - 1}{e^{-\theta} - 1}\right)$, where $\theta \in \Re$,

$$C_\theta(u, v) = -\frac{1}{\theta}\ln\left(1+\frac{\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}{\left(e^{-\theta}-1\right)}\right),$$

Kendall's $\tau = 1-\frac{4}{\theta}\left(1-D_1(\theta)\right)$, where $D_1(x) = \frac{1}{x}\int_0^x \frac{t}{e^t-1}dt$ is the Debye function.

2.4 Maximum Pseudolikelihood Method (MLE) of Copula Parameters Estimation

Suppose that a copula $C_\theta(u, v)$ is absolutely continuous with density

$$c_\theta(u,v) = \frac{\partial^2}{\partial u\,\partial v}C_\theta(u,v).$$

This method (described e.g. in [6]) involves maximizing a rank-based log-likelihood of this form

$$L(\theta) = \sum_{i=1}^{n}\ln\left(c_\theta\left(\frac{R_i}{n+1};\frac{S_i}{n+1}\right)\right), \tag{9}$$

where n is the sample size and θ is vector of parameters in the model. Arguments $\frac{R_i}{n+1},\frac{S_i}{n+1}$ equal to corresponding values of empirical marginal distributional functions of random variables X and Y.

This $L(\theta)$ function we use to define the *Akaike information criterion* (AIC) in the form (see e.g. [6])

$$AIC = -2L(\theta)+2k \tag{10}$$

where k is the number of independent parameters in the model.

AIC we use to compare the goodness of fit of our estimated model. A smaller AIC value means an improvement in the quality of the model fitting.

To obtain the initial values of the parameters for maximalization of the $L(\theta)$ function, we apply the *mean square error method*. It is based on the minimalization of the distance to the empirical copula

$$C_n(u,v) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\left(\frac{R_i}{n+1}\le u,\frac{S_i}{n+1}\le v\right).$$

2.5 Goodness of Fit Test for Copulas

Let $\{(x_j, y_j), j = 1, \ldots, n\}$ be n modeled 2-dimensional observations, F_X, F_Y their marginal distribution functions and F their joint distribution function.

We say that the class of copulas C_θ is correctly specified if there exists θ_0 so that

$$F(x, y) = C_{\theta_0}(F_X(x), F_Y(y))$$

holds.

White (1982) ([11]) showed that under correct specification of the copula class C_θ holds the following information matrix equivalence

$$-\boldsymbol{A}_{\theta_0} = \boldsymbol{B}_{\theta_0}$$

where

$$\begin{aligned} \boldsymbol{A}_\theta &= E\left[\nabla_\theta^2 \ln c_\theta(F_X(x), F_Y(y))\right] \\ \boldsymbol{B}_\theta &= E\left[\nabla_\theta \ln c_\theta(F_X(x), F_Y(y)) \nabla'_\theta \ln c_\theta(F_X(x), F_Y(y))\right] \end{aligned}$$

and c_θ is the density function of C_θ (copula C_θ must be absolutely continuous).

The testing procedure, which is proposed in [9], is based on the empirical distribution functions

$$\hat{F}_X(s) = \frac{1}{n}\sum_{i=1}^{n} 1(x_i \le s) \text{ and } \hat{F}_Y(s) = \frac{1}{n}\sum_{i=1}^{n} 1(y_i \le s)$$

and also on the consistent estimator $\hat{\theta}$ of θ_0 that maximizes $\sum_{i=1}^{n} \ln c_\theta\left(\hat{F}_X(x_i), \hat{F}_Y(y_i)\right)$.

To introduce the sample versions of $\boldsymbol{A}$ and $\boldsymbol{B}$ put

$$\begin{aligned} \boldsymbol{A_i}(\theta) &= \nabla_\theta^2 \ln c_\theta\left(\hat{F}_X(x_i), \hat{F}_Y(y_i)\right) \\ \boldsymbol{B_i}(\theta) &= \nabla_\theta \ln c_\theta\left(\hat{F}_X(x_i), \hat{F}_Y(y_i)\right) \nabla'_\theta \ln c_\theta\left(\hat{F}_X(x_i), \hat{F}_Y(y_i)\right) \end{aligned}$$

$$\hat{\boldsymbol{A}}_\theta = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{A}_i(\theta),$$

$$\hat{\boldsymbol{B}}_\theta = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{B}_i(\theta)$$

and

$$\boldsymbol{d}_i(\theta) = vech(\boldsymbol{A_i}(\theta) + \boldsymbol{B_i}(\theta)),$$

vech (***M***) is the vector of dimension *k* x 1 containing the upper triangle (in the lexicographic ordering) of the symmetric matrix **M** of the type *k* x *k* (where *k* is the dimension of the space of parameters θ).

Put $\hat{\boldsymbol{D}}_\theta = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{d}_i(\theta)$.

Under the hypothesis of proper specification the statistics $\sqrt{n}\hat{\mathbf{D}}_\theta$ has asymptotical distribution N(0, ***V***), where ***V*** is estimated by $\hat{\mathbf{V}} = \frac{1}{n-1}\sum \mathbf{d}_i'(\theta).\mathbf{d}_i(\theta)$.

Therefore

$$\chi^2 = n.\hat{\mathbf{D}}_\theta'.\hat{\mathbf{V}}_\theta^{-1}.\hat{\mathbf{D}}_\theta \tag{11}$$

is asymptotically as $\chi^2_{k(k+1)/2}$.

3 Results

In this section, we summarize all the results in tables and graphs. For our research we used 20 real data series (exchange rates, varied macroecomic data and other financial data series).

First, we 'fitted' these time series with the SETAR model (see [3]). We based the selection of the models (optimizing the number of states and the order of the local autoregressive models) on the BIC criterion (see, e.g. [1], [5]). Recall that the residuals of these models are supposed to be independent (not only serially non-correlated). This property can be tested by the BDS test (see [2]).

Inspired by the approach of Rakonczai (2009) ([10]), we applied autocopulas to the time series of the above-mentioned residuals in order to gauge how much they violate the assumptions of independence. If the test showed dependence in residuals, we described this dependence of time lagged residuals of SETAR models by means of copulas. For each couple $(\hat{e}_t, \hat{e}_{t-k})$ and each class of copulas we subsequently performed the following sequence of procedures:

a) calculation of ML estimates and AIC,

b) goodness of fit tests and corresponding p-values.

3.1 Results – The BDS Test

First we tested our real data series with the BDS test. Zero hypothesis is independence in residuals $\{\hat{e}_t\}$. We used significance level α = 0.05. In Table 1 we can see the results of the BDS test and the number of regimes of SETAR model for which it is used.

Table 1
Results of the BDS test

data	suitable for	***The BDS test***		
		p value (H_0: independent)		conclusion (α = 0.05)
		2 regimes	3 regimes	
HUF	3 regimes	0,039277	**0,491910**	*independent*
SKK	**linear**	**0,497770**		*independent*
PLN	3 regimes	0,021701	**0,062389**	*independent*
CZK	**linear**	**0,003660**		***dependent***
SVK unemploy	2 regimes	**0,129207**	0,028170	*independent*
SVK inflation	3 regimes	0,000371	**0,048212**	***dependent***
DoS USA	3 regimes	0,028691	**0,064259**	*independent*
GDP HUF	3 regimes	0,002610	**0,000015**	***dependent***
GDP SVK	3 regimes	0,016147	**0,029499**	***dependent***
GVA agri	3 regimes	0,154628	**0,489461**	*independent*
GVA constr	3 regimes	0,141906	**0,492069**	*independent*
GVA fin	3 regimes	0,024453	**0,490318**	*independent*
GVA industry	**linear**	**0,007034**		***dependent***
GVA other	3 regimes	0,022448	**0,493020**	*independent*
NofB10 SVK	3 regimes	0,011104	**0,048212**	***dependent***
NofB100 SVK	3 regimes	0,000190	**0,107229**	*independent*
CAP. GOODS	2 regimes	**0,013195**	0,051269	***dependent***
EMPLOY SVK	**linear**	**0,000081**		***dependent***
UNEMPLOY ocist	3 regimes	0,114267	**0,000244**	***dependent***
TRANSPORT SVK	3 regimes	0,153837	**0,211726**	*independent*

The BDS test determined dependence in residuals in 9 cases (from 20) and here we used the description of residual dependence with 'k-lag auto-copula'.

In the next section 3.2 we describe in detail the results for two time series. In the case of time series 'CZK' and 'GDP HUF', the independence is reached only for k = 23 (CZK) and k = 18 (GDP HUF); so for these time series the SETAR model is not appropriate and therefore results for this time series will not be mentioned. Results for all 7 remaining time series we will only present in the form of tables and graphs in section 3.3.

a) Unemployment (seasonally adjusted)

In case of the time series 'Unemployment (seasonally adjusted)' for 1-lagged residuals, the BDS test showed dependence. Therefore, we used the BDS test also for lag $k = 2,3,\ldots$ etc. to find out the couple (residuals and k-lag residuals) where the BDS test determines independence. In this case it is k = 9. For these time lagged residuals of the SETAR models, where we have dependence, we calculated Kendall τ. Then we described the dependence of the time lagged residuals of the SETAR models by means of an Archimedean class of copulas (Gumbel, strict Clayton and Frank). Then we tested the 'goodness' of the copulas with the Goodness of Fit test and finally we calculated the L2 norm distance and AIC to see which copula was the best for the description of our couples. All of these results are in *Table 2* and, for better illustration, these results are also in the graphs underneath.

Note: 'd' means *dependent* and 'i' *independent*

Table 2
Summarized results for the Kendall τ, parameters of copulas, GoF test, L2 norm and AIC in case of lag 1 to 9 for time series 'Unemployment (seasonally adjusted)'

lag		1	2	3	4	5	6	7	8	9
BDS test	**p value** (H_0 indep.)	$<10^{-6}$	0,00003	0,00036	0,00028	0,00018	0,00171	0,0014	0,00616	**0,0523**
	conclusion	d	d	d	d	d	d	d	d	i
Kendall tau		0,55206	0,4362	0,35148	0,32966	0,29495	0,32123	0,2694	0,2602	0,2462
parameters of copulas	Gumbel	2,13894	1,7221	1,52794	1,44275	1,39507	1,44965	1,0495	1,34303	
	Clayton	1,60655	1,0041	0,73268	0,55573	0,42898	0,43848	0,1329	0,43478	
	Frank	6,84353	4,6904	3,69323	3,21936	2,903	3,28623	0,3755	2,5776	
Good of fit test	Gumbel	0,13839	0,2802	0,34904	0,21064	0,11198	0,37077	0,3858	0,29188	
	Clayton	0,36567	0,1200	0,28685	0,45586	0,3514	0,43836	0,1724	0,04778	
	Frank	0,40747	0,321	0,33888	0,03204	0,44108	0,12063	0,2002	0,41271	
L2 norm distance	Gumbel	**0,87811**	**0,94**	**0,90618**	**1,19393**	**1,16001**	**1,40264**	**1,3079**	**0,99168**	
	Clayton	2,24868	2,4014	2,19219	2,6145	2,63743	3,07746	1,3501	1,93755	
	Frank	1,09679	1,2584	1,09337	1,31255	1,40455	1,66317	1,3125	1,16871	
AIC	Gumbel	-100,192	-57,399	-34,871	-27,957	-23,0306	-27,141	1,742	-17,275	
	Clayton	-80,2784	-42,485	-24,777	-14,898	-7,58769	-7,7959	1,5358	-8,3735	
	Frank	-97,7153	-56,299	-35,316	-28,278	-22,3284	-26,816	1,8188	-17,059	

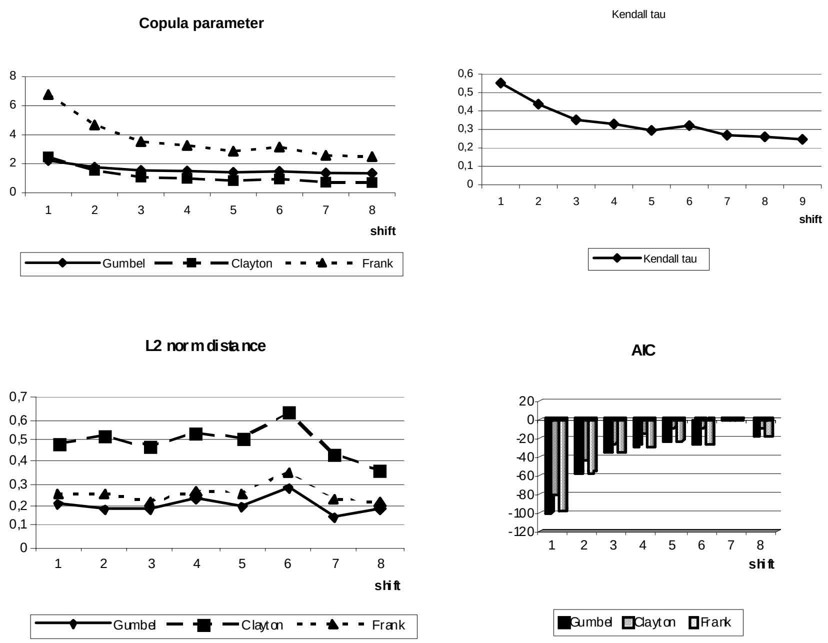

Figure 1
Graphs of parameters of copulas, Kendall τ , L2 norm and AIC in case of lag 1 to 9 for time series 'Unemployment (seasonally adjusted)'

For each couple $(\hat{e}_t, \hat{e}_{t-k})$, k = 1, …, 8, the optimal models in all three considered Archimedean copulas classes pass the GOF tests. The minimal values for the L2 norm was attained for the optimal model in the Gumbel class for all lag k = 1, …, 9. We observed that the autocopulas for the residuals were with increasing lag k near to the (independence indicating) product form. The value of Kendall τ also reduces with increasing lag k.

On the other side, because the independence is reached for high value *k* = 9, the SETAR model is not appropriate for these time series.

b) Inflation in Slovakia

In the case of the time series 'Inflation in Slovakia', the BDS test showed independence earlier, already for the lag 2, as we can see in Table 3 and Figure 2.

Table 3
Results for time series "Inflation in Slovakia"

lag	***BDS test***		**Kendall tau**	Good of fit test			copulas parameter		
	p value (H_0 independent)	conclusion		Gumbel	Clayton	Frank	Gumbel	Clayton	Frank
1	0,04821	d	0,2237	0,015551	0,48396	0,435379	1,26234	0,484523	2,19487
2	**0,05345**	i	0,0643	0,143151	0,40402	0,054031	1,11077	$8*10^{-6}$	0,58843

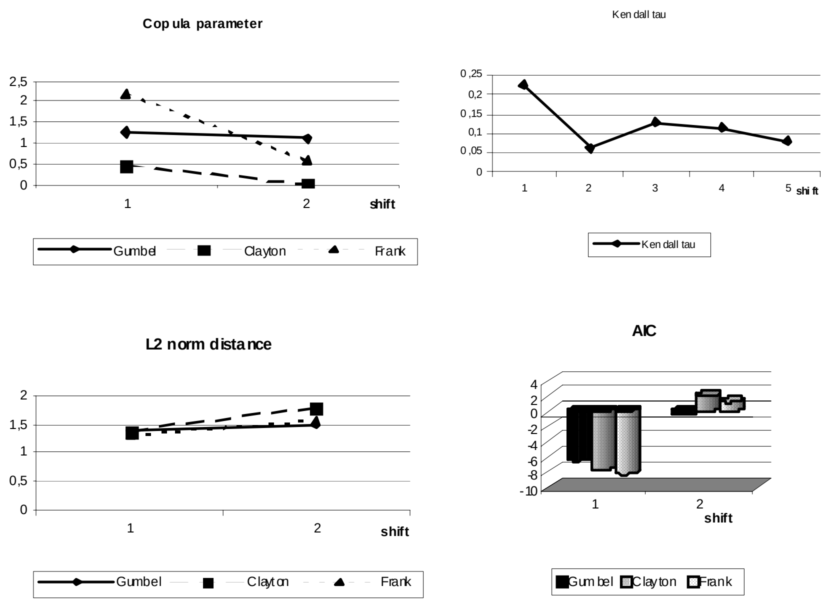

Figure 2
Graphs of parameters of copulas, Kendall τ , L2 norm and AIC in case of lag 1 and 2 for time series 'Inflation in Slovakia'

Among considered Archimedean copulas classes, only the Clayton and Frank class provide models (for k = 1) which were not subsequently rejected by the goodness of fit tests described above. The minimal values for the L2 norm and AIC was attained for the optimal model in the Frank class. The value of Kendall τ reduces with increasing lag *k*.

3.2 Results for Remaining Time Series in Graphs and Tables

In this section we can see results for all time series, where the BDS test in time lagged residuals ($k = 1$) rejected H_0 (except 'CZK' and 'GDP HUF') in tables and graphs.

a) p-value of BDS Test

In the next 7 pictures in Figure 3, we can see how the change p-value of BDS test until the residuals will be independent. We can see that for 4 time series the residuals are already independent for k = 2.

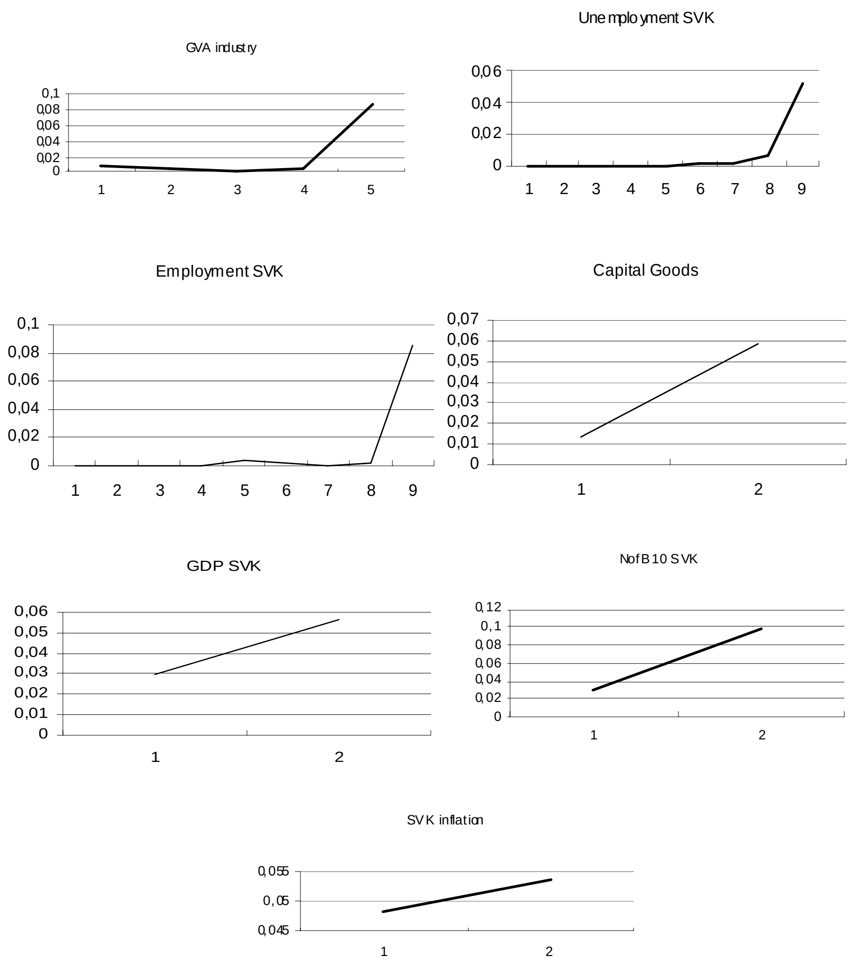

Figure 3
The graphs of changes of p-value of the BDS test

b) The Values of Kendall τ

In Figure 4 we can see that the growing lag k reduces the value of Kendall τ until the residuals are no longer dependent.

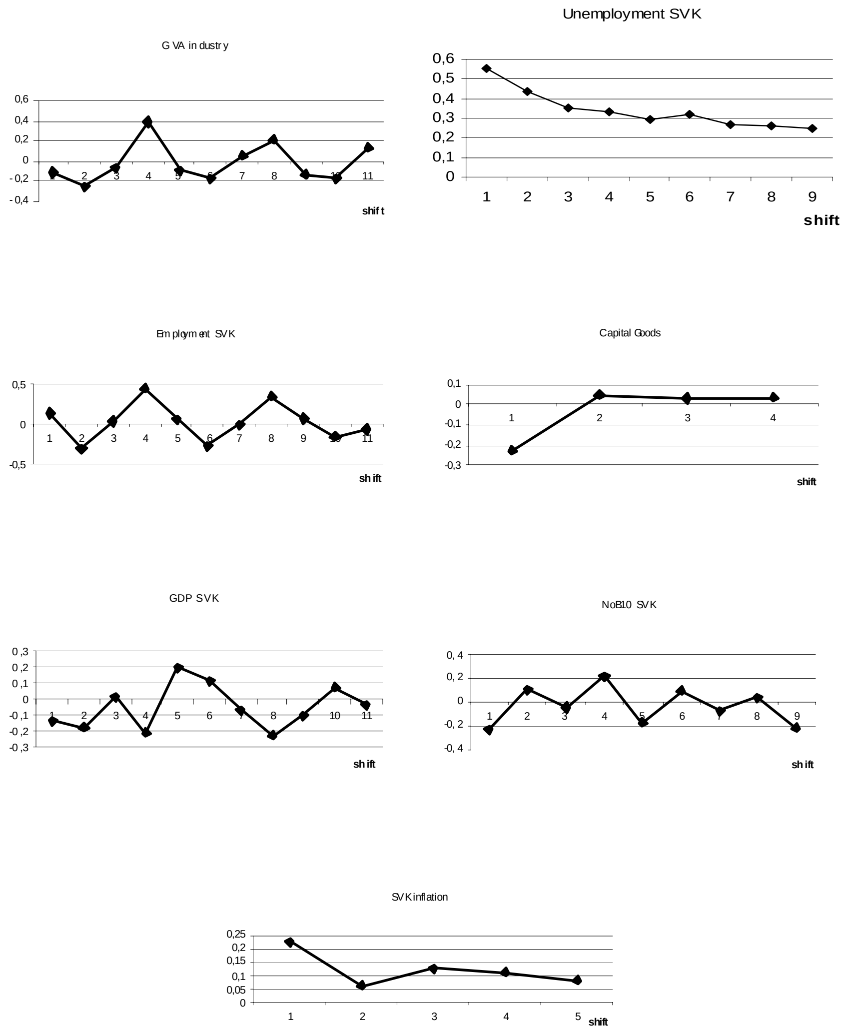

Figure 4
The graphs of changes of Kendall τ

c) Parameters of Copulas

The results for the parameters of the autocopulas are summarized in the Table 4.

Table 4

The table of changes of parameters of copulas when we aproach to the independence

Employ SVK	Gumbel	1,19905	1	1,04948	1,7498	1,07927	1,04948	1	1,04948
	Clayton	0,161757	0,1	0,132946	0,992168	1,00E-01	0,132946	0,058127	0,132946
	Frank	1,30708	-2,78766	0,375501	4,73045	0,586813	0,375501	0,005865	0,375501
Unemploy seasonal adjustment	Gumbel	2,13894	1,72208	1,52794	1,44275	1,39507	1,44965	1,04948	1,34303
	Clayton	1,60655	1,00458	0,732684	0,555728	0,428978	0,43848	0,132946	0,434783
	Frank	6,84353	4,69038	3,69323	3,21936	2,903	3,28623	0,375501	2,5776
GVA industry	Gumbel	1	1	1	1,57595	1			
	Clayton	1,00E-01	1,00E-01	1,00E-01	0,996258	1,00E-01			
	Frank	-1,02654	-2,29807	-0,672488	4,0506	-0,719123			
SVK inflation	Gumbel	1,26234	1,11077						
	Clayton	0,484523	8,01E-06						
	Frank	2,19487	0,588432						
GDP SVK	Gumbel	1	1						
	Clayton	1,00E-01	6,86E-05						
	Frank	-1,20478	-1,41876						
NofB10	Gumbel	1	1,18813						
	Clayton	1,00E-01	0,108703						
	Frank	-2,33575	0,913485						
Cap. Goods	Gumbel	1	1,0294						
	Clayton	0,1	1,00E-01						
	Frank	-2,28115	0,338111						

We can see how the parameters of the copulas change when we approach to independence. The parameter of the Gumbel copula approaches to 1, the parameter of the Clayton copula approaches to 0 and also the parameter of the Frank copula (in most cases) approaches to 0.

d) Goodness of Fit Test (GoF Test)

The results of the p-value of the GoF tests are summarized in the Table 5.

Table 5

The table of changes of p-values of GoF test

Employ SVK	Gumbel	0,49464	0,05061	0,38581	**0,00757**	0,39491	0,38581	0,20417	0,38581
	Clayton	0,25181	0.38163	0,17242	**0,01314**	0.30665	0,17242	0,28867	0,17242
	Frank	0,42371	0,24724	0,20022	0,40317	0,27459	0,20022	0,04971	0,20022

Unemploy seasonal adjustment	Gumbel	0,13839	0,28018	0,34904	0,21064	0,11198	0,37077	0,38581	0,29188
	Clayton	0,36567	0,12002	0,28685	0,45586	0,35140	0,43836	0,17242	**0,04778**
	Frank	0,40747	0,32097	0,33888	**0,03204**	0,44108	0,12063	0,20022	0,41271
GVA industry	Gumbel	0,30037	0,17217	0,36434	0,25722	0,49017			
	Clayton	0.14088	0.00253	0.32579	0,17110	0.05172			
	Frank	**0,03746**	0,18319	0,40797	0,17856	0,43309			
SVK inflation	Gumbel	**0,01555**	0,14315						
	Clayton	0,48396	0,40402						
	Frank	0,43538	0,05403						
GDP SVK	Gumbel	0,07065	0,09656						
	Clayton	0.31525	0,41367						
	Frank	0,06254	0,42859						
NofB10	Gumbel	0,23687	0,37940						
	Clayton	0.14667	0,18414						
	Frank	0,31182	0,34269						
Cap. Goods	Gumbel	0,12051	0,07906						
	Clayton	**0.02399**	0.35867						
	Frank	0,47009	0,46636						

Optimal values of the p-value (result of the GoF test) are bigger then 0.05 (a significant level) and in most cases it is fulfilled.

e) L2 Norm Distance

The values of L2 norm distance are in Table 6.

Table 6
The table of values of L2 norm distance

Employ SVK	Gumbel	**1,80711**	4,51264	**1,30795**	1,3504	**2,39809**	**1,30795**	**1,82147**	**1,30795**
	Clayton	2,27074	5.12645	1,35004	2,45897	2.66454	1,35004	1,84872	1,35004
	Frank	1,87764	**1,63538**	1,31251	**1,2516**	2,49996	1,31251	1,82138	1,31251
Unemploy seasonal adjustment	Gumbel	**0,878114**	**0,94**	**0,906184**	**1,19393**	**1,16001**	**1,40264**	**1,30795**	**0,991681**
	Clayton	2,24868	2,40135	2,19219	2,6145	2,63743	3,07746	1,35004	1,93755
	Frank	1,09679	1,25838	1,09337	1,31255	1,40455	1,66317	1,31251	1,16871
GVA industry	Gumbel	2,22829	3,81971	1,68091	1,65548	1,79534			
	Clayton	2.74804	4.46763	2.16497	1,86848	2.248			
	Frank	**1,57291**	**1,81806**	**1,45327**	**1,41446**	**1,57783**			
SVK inflation	Gumbel	1,36999	1,50498						
	Clayton	1,3878	1,79565						
	Frank	**1,3095**	1,55809						

GDP SVK	Gumbel	2,45152	3,07762						
	Clayton	2.96081	3,07802						
	Frank	**1,88178**	2,09258						
NofB10	Gumbel	4,24449	2,72947						
	Clayton	4.78234	2,88818						
	Frank	**2,86354**	2,75174						
Cap. Goods	Gumbel	4,14803	1,35069						
	Clayton	4.76407	1.43428						
	Frank	**1,96202**	1,33945						

We can see that for all four time series, in which the residuals are already independent for k = 2, the best copula is from the Frank class. In contrast, for time series in which the residuals are independent only for large values of lag k, the best copula is from the Gumbel class.

f) The Values of Information Criterion AIC

In the last section we can see in the table changes of AIC when we approach to independence.

Table 7
The table of changes of AIC when we aproach to the independence

Employ SVK	Gumbel	**-1,50845**	2	1,74197	-20,5521	**1,40842**	1,74197	2	1,74197
	Clayton	1,49185	4.5722	**1,5358**	-14,7951	2.26746	**1,5358**	**1,91957**	**1,5358**
	Frank	-0,170686	**-7,91672**	1,81884	**-21,1443**	1,56084	1,81884	1,99995	1,81884
Unemploy ocist	Gumbel	**-100,192**	**-57,3995**	-34,8711	-27,9567	**-23,0306**	**-27,1413**	1,74197	**-17,2749**
	Clayton	-80,2784	-42,4846	-24,7772	-14,8982	-7,58769	-7,79587	1,5358	-8,37347
	Frank	-97,7153	-56,2998	**-35,3158**	**-28,2783**	-22,3284	-26,8159	1,81884	-17,0593
GVA industry	Gumbel	2	2	2	-15,1784	2			
	Clayton	3.56993	5.221	3.18131	-16,3222	3.36876			
	Frank	**0,461541**	**-5,53732**	**1,3366**	**-17,7225**	**1,28324**			
SVK inflation	Gumbel	-6,50767	-0,487722						
	Clayton	-7,726	2,00002						
	Frank	**-8,30962**	1,19782						
GDP SVK	Gumbel	2	2						
	Clayton	3.40186	2,00121						
	Frank	**0,021092**	-0,586261						
NofB10	Gumbel	2	0,435091						
	Clayton	3.24797	1,9097						
	Frank	**-1,58858**	1,45354						
Cap. Goods	Gumbel	2	1,77872						
	Clayton	8.28015	3.19424						
	Frank	**-13,062**	1,64807						

From the table in this section we can see changes in the AIC values when we approach to independence. The smallest value of AIC (from our 3 families of copulas) means the best description of residuals. We see that in most cases the value of AIC confirms the findings of the value of L2 norms.

Conclusions

The topics of this paper were motivated by the modelling of a large number of economic and financial time series from emerging Central–European economies with the SETAR model (see [1], [5]).

We have based the selection of the models (optimizing the number of states and the order of the local autoregressive models) on the BIC criterion.

Recall that the residuals of these models are supposed to be independent (not only serially non-correlated). This property can be tested by e.g. the BDS test ([2]).

The BDS test has showed residual dependence (for $\alpha = 0.05$) in 9 cases (from 20) for the lag $k = 1$. We increase the lag k of residuals while they are independent. In the case of the time series 'CZK' and 'GDP HUF' the independence is reached only for k = 23 (CZK) and k = 18 (GDP HUF), so for these time series the SETAR model is not appropriate.

Inspired by the approach of Rakonczai (2009) [10] we applied autocopulas to the time series of the above-mentioned residuals in order to gauge how much they violate the assumptions of independence. We have arrived at an interesting conclusion concerning the residuals of the models that were selected as optimal on the basis of the BIC criterion. We have observed that the autocopulas for the residuals of the optimal models have been mostly substantially closer to the (independence indicating) product form (especially for lags k ≥ 2) than those for competing non-optimal models.

For all four time series, in which the residuals are already independent for k = 2, the best copula is from the Frank class. In contrast, for time series in which the residuals are independent only for large values of lag k, the best copula is from the Gumbel class.

In further research we would like to describe our time series with non-Archimedean copulas such as Gauss, Student copulas, etc. We would also like to use more complicated multi-regime models – for example the STAR and MSW models.

In this work we modeled residuals with bivariate copulas for couples $(\hat{e}_t, \hat{e}_{t-k})$, but we aim to model them with multivariate copulas for random vectors $(\hat{e}_t, \hat{e}_{t-1}, \cdots, \hat{e}_{t-k})$.

Acknowledgement

This work was supported by Slovak Research and Development Agency under contract No. LPP-0111-09.

References

[1] J. Arlt, M. Arltová: Financial Time Series (in Czech), Grada Publishing a.s., Praha, 2003

[2] W. A. Brock, W. D. Dechert, J. A. Scheinkman, B. Le Baron: A Test for Independence Based on the Correlation Dimension. Econometric Reviews 15, 1996, 197-235

[3] S. D. Campbell: Specification Testing and Semiparametric Estimation of Regime Switching Models: An Examination of the US Short Term Interest Rate. Department of Economics, Brown University, 2002

[4] P. Embrechts, F. Lindskog, McNeil, A.: Modeling Dependence with Copulas and Applications to Risk Management. In: Rachev, S. (Ed.) *Handbook of Heavy Tailed Distributions in Finance*. Elsevier, Chapter 8, 2001, pp. 329-384

[5] P. H. Frances, D. Van Dijk: Non-Linear Time Series Models in Empirical Finance, Cambridge University Press, Cambridge, 2000

[6] C. Genest, A. C. Favre: Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask. *Journal of Hydrologic Engineering* July/August 2007, pp. 347-368

[7] H. Joe: Multivariate Models and Dependence Concepts. Chapman & Hall, London, UK, 1997

[8] R. B. Nelsen: An Introduction to Copulas. Springer, New York, 1999

[9] A. Prokhorov: A Goodness-of-Fit Test for Copulas. MPRA Paper No. 9998, online at http://mpra.ub.uni-muenchen.de/9998, 2008

[10] P. Rakonczai: On Modeling and Prediction of Multivariate Extremes. Mathematical Statistics Centre for Mathematical Sciences, Lund University, 2009

[11] H. White: Maximum Likelihood Estimation of Misspecified Models. *Econometrica* **50**, 1982, pp. 1-26

[12] Magda Komorníková, Jozef Komorník: A Copula-based Approach to the Analysis of the Returns of Exchange Rates to EUR of the Visegrád Countries, in Acta Polytechnica Hungarica, Vol. 7, No. 3, 2010, pp. 79-91